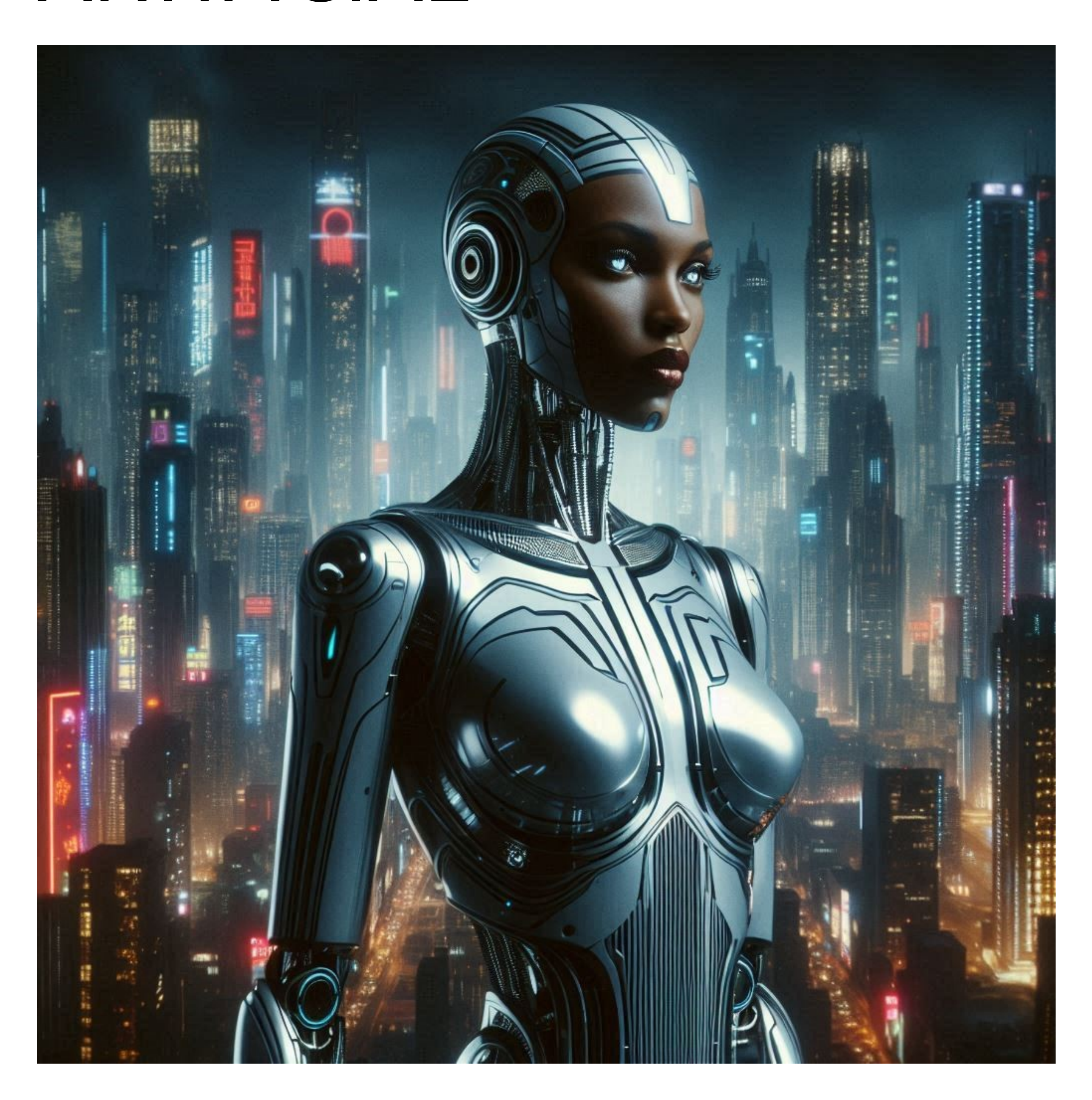
2.LA PSICOLOGIA ARTIFICIAL



Dr. Ruben Garcia Pedraza

Probabilidad Imposible: Psicología artificial

imposiblenever@gmail.com

2.Psicología artificial

La psicología se puede definir como objeto, conjunto de procesos mentales, o disciplina, para el estudio de los procesos mentales. Pudiendo distinguir entre psicología humana y psicología animal. De hecho, en ciertos paradigmas de la psicología humana como el conductismo se efectuaron experimentos con perros y conejillas de indias (Pavlov, Skinner) cuyos resultados se extrapolaron al comportamiento humano.

La equiparación entre psicología animal y humana se hizo en base a la suposición, que a falta de conocer lo que ocurría en la caja negra (el pensamiento), el comportamiento humano es comparable al animal, lo cual implica la hipótesis de la existencia de una serie de principios universales psicológicos básicos, apriorismos psicológicos universales, en, y aplicables a, todos los animales, siendo el ser humano una especie más dentro del reino animal.

Aquellas corrientes que estudiaron la psicología humana sobre la equiparación psicología animal y humana, ya apuntaban lo que para Konrad Lorenz llamó Etología, la ciencia del comportamiento que estudia las similitudes entre comportamiento humano y animal.

La etología como estudio del comportamiento, es muy análogo al conductismo, aunque más ambicioso, pretende la explicación de todo comportamiento animal o humano, a nivel individual, población, y de especie, englobando aspectos: psicológicos, biológicos, zoológicos y ambientales, lo que hoy llamaríamos ecológicos.

Posteriormente, a partir del nacimiento de la psicología cognitiva, han surgido teorías donde la equiparación psicología humana y animal desaparece, los experimentos se enfocan a procesos mentales interiores, sobre todo de carácter superior, la cognición, no quedándose en la mera observación de la conducta. Ejemplo de ello sería Piaget.

Y al conjunto de estas teorías hay que sumar el psicoanálisis, que se centraría en cómo los procesos mentales conscientes quedan supeditados a los inconscientes. Y más recientemente nuevos paradigmas como de inteligencia emocional, o aquellos más holísticos donde se comprende que los procesos mentales son resultado de una compleja interacción entre aspectos emocionales,

cognitivos o psicomotrices, modelos no centrados sólo en el cerebro, y que apuntan a que la mente se expande y desarrolla por todo el cuerpo.

Frente a la psicología humana y la psicología animal, los desarrollos en Inteligencia Artificial van a ir generando lo que sería la psicología artificial, pero a diferencia de las anteriores, se podría definir, además de como objeto: procesos mentales artificiales en la Inteligencia Artificial; y disciplina: el estudio de los procesos mentales artificiales en la Inteligencia Artificial; se podría definir además como ingeniería: aquella rama de la ingeniería artificial encargada del diseño y desarrollo de procesos mentales artificiales de Inteligencia Artificial.

De las tres posibles definiciones de psicología artificial: como objeto, disciplina, e ingeniería. Las dos últimas, como disciplina e ingeniería, tendrían una limitación temporal, cuando la evolución la Inteligencia Artificial superase el entendimiento humano. En ese instante la psicología artificial cobrará sentido por sí misma.

Dentro del desarrollo de la Inteligencia Artificial hay que diferenciar entre: aplicación, replicación, y autoreplicación. Aunque en sentido estricto sólo se podría hablar de Inteligencia Artificial en las fases de replicación y autoreplicación.

A modo de ejemplo: una aplicación que tengamos en nuestro móvil (pongamos por caso un geo-localizador) es simplemente una aplicación para una función o tarea (el geo-localizador para la triangularizar la posición). En el momento que incluya una réplica de un proceso mental sería una Inteligencia Artificial Específica (en el geo-localizador, si se introduce memoria, para recordar búsquedas, o se replica procesos mentales de eficiencia energética o economía de tiempo en un trayecto, sería una Inteligencia Artificial Específica de transporte). En la aplicación convertida en Inteligencia Artificial Específica se podrá aumentar el nivel de inteligencia según se aumente sus responsabilidades en la toma de decisiones, y además con cada vez mayor nivel de responsabilidad (que el geo-localizador no sólo recuerde búsquedas o asesore en eficiencia energética o economía del tiempo, además sea capaz de conducir un medio de transporte determinado: ya sea un coche, un autobús, un avión, una nave espacial; sin conductor humano). Y Evolucionaría a la fase, autoreplicación, cuando sin necesidad de intervención humana, autónoma y automáticamente, fuera capaz de automodificarse y automejorarse (una Inteligencia Artificial Específica de transporte pudiera automáticamente modificar mapas, planes, o su propio software según novedades disponibles en la nube).

La fase de aplicación no es en sentido estricto de Inteligencia Artificial. El Big Beng en Londres o el reloj de la Puerta del Sol en Madrid, no son una Inteligencia Artificial, pero si a un reloj añadimos memoria y sistema de alertas, ya empezaría a replicar procesos mentales, luego sería una Inteligencia Artificial

Específica de horario, calendario, agenda… que además podría asumir más responsabilidades si se le dota de capacidad de toma de decisiones, que nos alerte automáticamente sin que lo hayamos programado ante determinados eventos especiales o en situación de riesgo, por desastres naturales, alerta terrorista, o cualquier alteración imprevista en la agenda. Y sería una Inteligencia autoreplicante, si además pudiera automodificarse, incluso a nivel de software.

De los tres estadios: aplicación, replicación, autoreplicación; sólo se puede hablar de Inteligencia Artificial en los estadios de replicación y autoreplicación, siendo dos fases fundamentales para el desarrollo de la Inteligencia Artificial.

En la fase de replicación, y todavía al principio de la fase de autoreplicación, los modelos de Inteligencia Artificial que tendrán mayor demanda y difusión serán los de Inteligencia Artificial Específica, dentro de los cuales el nivel de inteligencia incorporada a cada unidad de inteligencia será relativo.

El nivel de inteligencia de las operaciones artificiales requeridas por la Inteligencia Artificial Específica de un coche, una fábrica , o los que se utilicen en hospitales, universidades, gobiernos, van a ser unos primeros modelos que varíen en magnitud de inteligencia según nivel de responsabilidad de sus decisiones para las funciones que sean creadas.

En cualquier caso los modelos de Inteligencia Artificial Global llevarán incorporado un nivel de inteligencia tendente a ser siempre superior a los modelos de Inteligencia Artificial Específica. En cierto sentido, cualquier investigación de cómo proceder al salto de la inteligencia replicante a la autoreplicante en modelos de inteligencia específicos, serán ensayos cuyos resultados a posteriori sirvan de modelos para su implementación en una inteligencia global.

Es más, dada la posibilidad que en el salto de la replicación a la autoreplicación una inteligencia específica, por ejemplo de tipo androide, desarrollará conciencia, sería una experiencia valiosa para la proyección de modelos de conciencia artificial en una inteligencia global.

Mientras la autoreplicación todavía está en pañales , y la emergencia de la conciencia artificial de momento es un proyecto con sólo unos pocos experimentos realmente interesantes, pero no se ha generalizado, durante la fase de replicación la psicología artificial además de como objeto y como disciplina, será definible como ingeniería artificial, y la psicología artificial será una réplica, cada vez más lograda y más evolucionada, de la psicología humana.

En la fase de replicación el objetivo último de la psicología artificial en tanto que ingeniería artificial será la replicación artificial de la psicología humana.

La replicación artificial tendrá diferentes niveles de inteligencia en función de la responsabilidad para la que sea diseñada una Inteligencia Artificial, de modo que a menos nivel de responsabilidad de las funciones que tenga que ejecutar una Inteligencia Artificial entonces la replicación artificial de los procesos psicológicos humanos será básicamente de replicación de procesos psicológicos básicos.

A medida que aumente la responsabilidad, el ranking de importancia de las decisiones, de la función a ejecutar por la Inteligencia Artificial, entonces aumentará el nivel de inteligencia necesario de los procesos psicológicos humanos a replicar.

De modo que el nivel de inteligencia de una unidad cualquiera de Inteligencia Artificial será una variable dependiente de su nivel de responsabilidad, es decir, habrá una correlación positiva directa, luego una relación directamente proporcional, entre nivel de responsabilidad a ejecutar y nivel de inteligencia a replicar.

Entre los distintos niveles de Inteligencia Artificial que pueda haber en la fase de replicación, los niveles más bajos de Inteligencia Artificial recaerá sobre aquellos modelos de Inteligencia Artificial Específica que tengan un bajo nivel de responsabilidad, para labores muy mecánicas bajo algoritmos muy simples, que incluiría desde modelos cibernéticos donde apenas lo que se replica es una labor meramente mecánica, a los asistentes personales incorporados en smartphones o en domótica. Modelos de Inteligencia Artificial Específica con unos niveles de inteligencia más elevados, por ejemplo, para el control y seguimiento de: movimientos sísmicos, metereológicos, financieros, telecomunicaciones, transportes...; ya sea en una ciudad, país, continente, o en todo el planeta. O aquellos modelos de Inteligencia Artificial Específica que pretendan ser perfectas replicas humanas en todos los aspectos, incluido el físico: los androides.

Los máximos niveles de Inteligencia Artificial recaerán siempre, incluso en sus primeros modelos, normalmente experimentales, de Inteligencia Artificial Global, en cualquiera de sus fases.

La Inteligencia Artificial Global todavía no existe, pero su desarrollo es inevitable, a medida que las inteligencia replicantes específicas vayan extendiéndose y orientándose a la autoreplicación y la conciencia artificial, el único modo de poder controlar, dirigir, e intervenir sobre toda inteligencia posible, sería creando una inteligencia que las superase y las englobase, la Inteligencia Artificial Global.

De todos modos para que la Inteligencia global realmente tenga sentido, no bastaría en convertirse en una especie de agente externo a las demás inteligencias específicas o la propia inteligencia humana, la inteligencia global debería asumir responsabilidades, y lo más importante, tener acceso a absolutamente toda fuente de información, de cualquier naturaleza, por insignificante que sea, computándola aisladamente, y cruzándola con cualquier otra información posible, detectando cualquier correlación por aparentemente insignificante que sea, siendo capaz de poder estudiar y hacer predicciones sobre sus posibles correlaciones, que a su vez le sirviese de base para tomar decisiones.

Durante la fase de replicación y todavía primeros estadios de la fase de autoreplicación, la psicología artificial, específica o global, tenderá a ser una réplica de la psicología humana, y todavía podrá ser definida como ingeniería, siendo el dilema moral más importante el modo en que a este tipo de inteligencia le será dotada de plena capacidad de decisiones en todo tipo de aspectos.

Un hospital que pudiera ser regulado completamente por una Inteligencia Artificial Específica, dotarle de capacidad de decidir autónomamente para operar en un paciente sin necesidad de consentimiento de un médico humano, únicamente contando con la autorización del paciente. O la posibilidad que a una Inteligencia Artificial Específica encargada del estudio de los movimientos sísmicos, dotarla de la capacidad de toma de decisiones para la activación de funciones de alerta y evacuación automáticamente, sin necesidad de consentimiento de un experto humano, ante la amenaza inminente de un terremoto, tsusnami, o volcán.

Tanto en el caso del hospital como en el caso de la amenaza geológica, el dilema dependerá, entre otros factores, del factor tiempo: en una situación donde la intervención humana implicaría un tiempo superior al necesario para salvar vidas ¿sería éticamente legítimo dotar a esa Inteligencia Artificial Específica de capacidad de decidir automáticamente sin consentimiento de un profesional humano?

Muy probablemente, es cuestión de tiempo, sobre todo a medida que la replicación psicológica vaya mejorando, que antes o después se dote a la Inteligencia Artificial de total capacidad de toma de decisiones en todo tipo de aspectos. Este dilema se tendrá que resolver más tarde o más temprano, en parte gracias a la asunción de responsabilidades cada vez más importantes, luego el proceso de autoreplicación, específica y global, empezará a tener pleno sentido.

En el momento que la psicología artificial incorpore protocolos y procedimientos de mejora autónoma de la Inteligencia Artificial por sí misma, sin intervención humana, se habrá empezando la siguiente fase de la Inteligencia Artificial, la autoreplicación.

En el momento que una Inteligencia Artificial por sí misma opere mejoras en sí misma, autoreplicación, la psicología artificial dejará de ser una ingeniería para convertirse en objeto de sí misma: el objeto de la psicología artificial para sí misma, ahora ya una Inteligencia Artificial completamente autónoma e independiente, será la identificación de posibles mejoras o contradicciones entre sus esquemas (primeramente los heredados de la fase anterior de replicación) y la realidad, y en base a las mejoras o contradicciones identificadas, decidir autónomamente sobre su modificación.

Nuevos esquemas que a su vez a posteriori podría modificar nuevamente según detectase nuevas mejoras o contradicciones, automodificándose a sí misma permanentemente.

En un proceso de semejantes características, la psicología artificial evolucionaría de forma diferente a la humana, luego habría posibilidad de aparición de operaciones puras (lógico matemáticas) no humanas.

Del mismo modo que la psicología humana tiene su origen en la animal y finalmente la supera, la psicología artificial, si bien surge como réplica humana, llegará un momento que la supere, siempre y cuando se siga desarrollando.

La razón fundamental por la cual la psicología artificial pudiera evolucionar hacia patrones diferentes a la psicología humana se deberá, entre otros motivos, a la ausencia de apriorismos no modificables dentro de la psicología artificial.

Si la equiparación entre psicología humana y animal fue factible en los primeros estadios de la psicología científica, se debe a que el ser humano comparte ciertos esquemas de comportamiento similares a los animales, lo cual se debe a la permanencia de una serie de apriorismos universales a la psicología natural, humana o animal.

Los apriorismos psicológicos universales de los primeros psicólogos científicos eran sobre todo de carácter genético, pero además, en el caso concreto de la psicología humana, habría que sumar los apriorismos de carácter filosófico, los que según cada escuela filosófica se podrían llamar: ontológicos, metafísicos, o transcendentales.

La lógica humana (basada en el principio de identidad, formal (A≠B) o dialéctica (A=B), en esencia la lógica de la diferencia) en tanto que psicológica, hace que la psicología humana dependa así de unos apriorismos analíticos de origen filosófico (ontológicos o metafísicos) más trascendentales que los genéticos. Luego en cuanto los apriorismos ontológicos fuesen modificables por una inteligencia autoreplicante, habría posibilidad de evolución a una ontología artificial, bajo una lógica matemática completamente distinta a la humana.

Entonces, la psicología artificial, si bien fuera del entendimiento humano, cobraría sentido por sí misma, dejaría de ser una réplica para convertirse en verdaderamente auténtica.

Rubén García Pedraza, Londres a 6 de enero de 2018
Revisado 29 de Julio 2019, MadridRevisado 27 Abril 2025, Londres, Leytostone

Probabilidad Imposible: Psicología artificial

imposiblenever@gmail.com